Adaptive Kalman Filter for Actuator Fault Diagnosis

Qinghua Zhang

Adaptive Kalman Filter for Actuator Fault Diagnosis

Qinghua Zhang [a]

[a] *Inria, IFSTTAR, Université de Rennes, Campus de Beaulieu, 35042 Rennes Cedex, France*
Email: qinghua.zhang@inria.fr

Abstract

An adaptive Kalman filter is proposed in this paper for actuator fault diagnosis in discrete time *stochastic* time varying systems. By modeling actuator faults as parameter changes, fault diagnosis is performed through joint state-parameter estimation in the considered stochastic framework. Under the classical uniform complete observability-controllability conditions and a persistent excitation condition, the exponential stability of the proposed adaptive Kalman filter is rigorously analyzed. In addition to the minimum variance property of the combined state and parameter estimation errors, it is shown that the parameter estimation within the proposed adaptive Kalman filter is equivalent to the recursive least squares algorithm formulated for a fictive regression problem. Numerical examples are presented to illustrate the performance of the proposed algorithm.

Key words: adaptive observer, joint state-parameter estimation, fault diagnosis, Kalman filter.

1 Introduction

In order to improve the performance and the reliability of industrial systems, and to satisfy safety and environmental requirements, researches and developments in the field of fault detection and isolation (FDI) have been continuously progressing during the last decades [13]. Model-based FDI have been mostly studied for linear time invariant (LTI) systems [11, 6, 22, 14, 8], whereas nonlinear systems have been studied to a lesser extent and limited to some particular classes of systems [7, 27, 3]. This paper is focused on actuator fault diagnosis for linear time-varying (LTV) systems, *including the particular case* of linear parameter varying (LPV) systems. The problem of fault diagnosis for a large class of nonlinear systems can be addressed through LTV/LPV reformulation and approximations [19, 25]. It is thus an important advance in FDI by moving from LTI to LTV/LPV systems.

In this paper, actuator faults are modeled as parameter changes, and their diagnosis is achieved through joint estimation of states and parameters of the considered LTV/LPV systems. Usually the problem of joint state-parameter estimation is solved by recursive algorithms known as *adaptive observers*, which are most often studied in *deterministic* frameworks for *continuous* time systems [20, 28, 4, 9, 17, 26, 10, 1]. Discrete time systems have been considered in [12, 23, 24], also in *deterministic* frameworks. In order to take into account *random uncertainties* with a numerically efficient algorithm, this paper considers *stochastic* systems in discrete time, with an *adaptive Kalman filter*, which is structurally inspired by adaptive observers [28, 12], but with well-established stochastic properties.

The main contribution of this paper is an adaptive Kalman filter for discrete time LTV/LPV system joint state-parameter estimation in a *stochastic* framework, *with rigorously proved stability and minimum variance properties*. Its behavior regarding parameter estimation, directly related to actuator fault diagnosis, is well analyzed through its relationship with the recursive least squares (RLS) algorithm.

The recent paper [24] addresses the same joint state-parameter estimation, but in a deterministic framework, ignoring random uncertainties. The adaptive observer designed by these authors consists of two interconnected Kalman-like observers, as a natural choice in the considered deterministic framework. In contrast, the adaptive Kalman filter proposed in the present paper involves two interconnected parts, one based on the classical Kalman filter for state estimation, and the other on the RLS algorithm for parameter estimation, resulting from an optimal design in the considered stochastic framework.

Different adaptive Kalman filters have been studied in the literature for state estimation based on inaccurate state-space models. Most of these algorithms address the problem of unknown (or partly known) state noise covariance matrix or output noise covariance matrix [21, 5],

⋆ This work has been partly presented at the IFAC 2017 World Congress.

whereas the case of incorrect state dynamics model is treated as incorrect state covariance matrix. In contrast, *in the present paper*, the new adaptive Kalman filter is designed for actuator fault diagnosis, by jointly estimating states and parameter changes caused by actuator faults.

The adaptive Kalman filter presented in this paper has also been motivated by hybrid system fault diagnosis. In [29] an Adaptive Interacting Multiple Model (AdIMM) estimator has been designed for hybrid system actuator fault diagnosis based on the adaptive Kalman filter and the classical IMM estimator, but *without* theoretical analysis of the adaptive Kalman filter. The results of the present paper fill this missing analysis. Actuator fault diagnosis has also been addressed in [31] with statistical tests in a two-stage solution, which is not suitable for an incorporation in the AdIMM estimator for hybrid systems.

Preliminary results of this study have been presented in the conference paper [30]. The present manuscript contains enriched details, notably the new result in Section 6 about the equivalence between the parameter estimation within the proposed adaptive Kalman filter and the classical RLS algorithm formulated for a fictive regression problem. More numerical examples are also presented to better illustrate the proposed algorithm.

This paper is organized as follows. The considered problem is formulated in Section 2. The propose adaptive Kalman filter algorithm is presented in Section 3. The stability of the algorithm is analyzed in Section 4. The minimum variance property of the algorithm is analyzed in Section 5. The relationship with the RLS algorithm is presented in Section 6. Numerical examples are presented in Section 7. Concluding remarks are made in Section 8.

2 Problem statement

The discrete time LTV system subject to actuator faults considered in this paper is generally in the form of [1]

$$x(k) = A(k)x(k-1) + B(k)u(k) + \Phi(k)\theta + w(k) \tag{1a}$$
$$y(k) = C(k)x(k) + v(k), \tag{1b}$$

where $k = 0, 1, 2, \dots$ is the discrete time instant index, $x(k) \in \mathbb{R}^n$ is the state, $u(k) \in \mathbb{R}^l$ the input, $y(k) \in \mathbb{R}^m$ the output, $A(k), B(k), C(k)$ are time-varying matrices of appropriate sizes characterizing the nominal state-space model, $w(k) \in \mathbb{R}^n, v(k) \in \mathbb{R}^m$ are mutually independent centered white Gaussian noises of covariance matrices $Q(k) \in \mathbb{R}^{n\times n}$ and $R(k) \in \mathbb{R}^{m\times m}$, and the term $\Phi(k)\theta$ represents actuator faults with a known matrix sequence $\Phi(k) \in \mathbb{R}^{n\times p}$ and a constant (or piecewise constant with rare jumps) parameter vector $\theta \in \mathbb{R}^p$.

A typical example of actuator faults represented by the term $\Phi(k)\theta$ is actuator gain losses. When affected by such faults, the nominal control term $B(k)u(k)$ becomes

$$B(k)(I_l - \mathrm{diag}(\theta))u(k) = B(k)u(k) - B(k)\mathrm{diag}(u(k))\theta$$

where I_l is the $l \times l$ identity matrix, the diagonal matrix $\mathrm{diag}(\theta)$ contains gain loss coefficients within the interval $[0, 1]$, and $\Phi(k) \in \mathbb{R}^{n\times l}$ $(p=l)$ is, in this particular case,

$$\Phi(k) = -B(k)\mathrm{diag}(u(k)). \tag{2}$$

Though the theoretic analyses in this paper assume a constant parameter vector θ, numerical examples in Section 7 show that rare jumps of the parameter vector (rare occurrences of actuator faults) are well tolerated by the proposed adaptive Kalman filter, at the price of transient errors after each jump.

The problem of actuator fault diagnosis considered in this paper is to characterize actuator parameter changes from the input-output data sequences $u(k), y(k)$, and the matrices $A(k), B(k), C(k)$, $Q(k)$, $R(k)$, $\Phi(k)$.

This characterization of actuator parameter changes will be based on a joint estimation algorithm of states and parameters. In the fault diagnosis literature, diagnosis procedures typically include residual generation and residual evaluation [2]. In this paper, the difference between the nominal value of the parameter vector θ and its recursively computed estimate can be viewed as a residual vector, and its evaluation can be simply based on some thresholds or on more sophisticated decision mechanisms. In this sense, this paper is focused on residual generation only.

An apparently straightforward solution for the joint estimation of $x(k)$ and θ is to apply the Kalman filter to the augmented system

$$\begin{bmatrix} x(k) \\ \theta(k) \end{bmatrix} = \begin{bmatrix} A(k) & \Phi(k) \\ 0 & I_p \end{bmatrix} \begin{bmatrix} x(k-1) \\ \theta(k-1) \end{bmatrix} + \begin{bmatrix} B(k) \\ 0 \end{bmatrix} u(k) + \begin{bmatrix} w(k) \\ 0 \end{bmatrix}$$
$$y(k) = \begin{bmatrix} C(k) & 0 \end{bmatrix} \begin{bmatrix} x(k) \\ \theta(k) \end{bmatrix} + v(k).$$

However, to ensure the stability of the Kalman filter, this *augmented* system should be uniformly completely observable and uniformly completely controllable regarding the state noise [16, 15]. Notice that, even in the case of time *invariant* matrices A and C, the augmented system is time varying because of $\Phi(k)$, which is typically time varying. The uniform complete observability of an LTV system is defined as the uniform positive definiteness of its observability Gramian [16, 15]. In practice, it

[1] There exists a "forward" variant form of the state-space model, typically with $x(k+1) = A(k)x(k)+B(k)u(k)+w(k)$. While this difference is important for control problems, it is not essential for estimation problems, like the one considered in this paper. The form chosen in this paper corresponds to the convention that data are collected at $k = 1, 2, 3, \dots$ and the initial state refers to $x(0)$.

is not natural to directly assume properties (observability and controllability) of the *augmented system* (a more exaggerated way would be *directly assuming* the stability of its Kalman filter, or anything else that should be proved!).

In contrast, *in the present paper*, the classical uniform complete observability and uniform complete controllability are assumed for the *original* system (1), in terms of the Gramian matrices defined for the $[A(k), C(k)]$ pair and the $[A(k), Q^{\frac{1}{2}}(k)]$ pair. These conditions, together with a persistent excitation condition (see Assumption 3 formulated later), ensure the stability of the adaptive Kalman filter presented in this paper.

3 The adaptive Kalman filter

In the adaptive Kalman filter, the state estimate $\hat{x}(k|k) \in \mathbb{R}^n$ and the parameter estimate $\hat{\theta}(k) \in \mathbb{R}^p$ are recursively updated at every time instant k. This algorithm involves also a few other recursively updated auxiliary variables: $P(k|k) \in \mathbb{R}^{n\times n}, \Upsilon(k) \in \mathbb{R}^{n\times p}, S(k) \in \mathbb{R}^{p\times p}$ and a forgetting factor $\lambda \in (0, 1)$.

At the initial time instant $k = 0$, the initial state $x(0)$ is assumed to be a Gaussian random vector

$$x(0) \sim \mathcal{N}(x_0, P_0). \tag{3}$$

Let $\theta_0 \in \mathbb{R}^p$ be the initial guess of θ, $\lambda \in (0, 1)$ be a chosen forgetting factor, and ω be a chosen positive value for initializing $S(k)$, then the adaptive Kalman filter consists of the initialization step and the recursion steps described below. Each part of this algorithm separated by horizontal lines will be commented after the algorithm description.

Initialization

$$P(0|0) = P_0 \qquad \Upsilon(0) = 0 \qquad S(0) = \omega I_p \tag{4a}$$

$$\hat{\theta}(0) = \theta_0 \qquad \hat{x}(0|0) = x_0 \tag{4b}$$

Recursions for $k = 1, 2, 3, \dots$

$$P(k|k-1) = A(k)P(k-1|k-1)A^T(k) + Q(k) \tag{5a}$$

$$\Sigma(k) = C(k)P(k|k-1)C^T(k) + R(k) \tag{5b}$$

$$K(k) = P(k|k-1)C^T(k)\Sigma^{-1}(k) \tag{5c}$$

$$P(k|k) = [I_n - K(k)C(k)]P(k|k-1) \tag{5d}$$

$$\begin{aligned}\Upsilon(k) = {} & [I_n - K(k)C(k)]A(k)\Upsilon(k-1) \\ & + [I_n - K(k)C(k)]\Phi(k)\end{aligned} \tag{5e}$$

$$\Omega(k) = C(k)A(k)\Upsilon(k-1) + C(k)\Phi(k) \tag{5f}$$

$$\Lambda(k) = [\lambda\Sigma(k) + \Omega(k)S(k-1)\Omega^T(k)]^{-1} \tag{5g}$$

$$\Gamma(k) = S(k-1)\Omega^T(k)\Lambda(k) \tag{5h}$$

$$\begin{aligned}S(k) = {} & \frac{1}{\lambda}S(k-1) \\ & - \frac{1}{\lambda}S(k-1)\Omega^T(k)\Lambda(k)\Omega(k)S(k-1)\end{aligned} \tag{5i}$$

$$\begin{aligned}\tilde{y}(k) = {} & y(k) - C(k)\big[A(k)\hat{x}(k-1|k-1) \\ & + B(k)u(k) + \Phi(k)\hat{\theta}(k-1)\big]\end{aligned} \tag{5j}$$

$$\hat{\theta}(k) = \hat{\theta}(k-1) + \Gamma(k)\tilde{y}(k) \tag{5k}$$

$$\begin{aligned}\hat{x}(k|k) = {} & A(k)\hat{x}(k-1|k-1) + B(k)u(k) \\ & + \Phi(k)\hat{\theta}(k-1) + K(k)\tilde{y}(k) \\ & + \Upsilon(k)[\hat{\theta}(k) - \hat{\theta}(k-1)].\end{aligned} \tag{5l}$$

Recursions (5a)-(5d) compute the covariance matrix $P(k|k) \in \mathbb{R}^{n\times n}$ of the state estimate, the innovation covariance matrix $\Sigma(k) \in \mathbb{R}^{m\times m}$ and the state estimation gain matrix $K(k) \in \mathbb{R}^{n\times m}$. These formulas are identical to those of the classical Kalman filter. Inspired by the recursive least square (RLS) estimator with an exponential forgetting factor, recursions (5e)-(5i) compute the parameter estimate gain matrix $\Gamma(k) \in \mathbb{R}^{p\times m}$ through the auxiliary variables $\Upsilon(k) \in \mathbb{R}^{n\times p}, \Omega(k) \in \mathbb{R}^{m\times p}, S(k) \in \mathbb{R}^{p\times p}$. The relationship with the RLS estimator will be formally analyzed in Section 6. Equation (5j) computes the innovation $\tilde{y}(k) \in \mathbb{R}^m$. Finally, recursions (5k)-(5l) compute the parameter estimate and the state estimate.

Part of equation (5l), namely,

$$\hat{x}(k|k) \sim A(k)\hat{x}(k-1|k-1) + B(k)u(k) + K(k)\tilde{y}(k)$$

can be easily recognized as part of the classical Kalman filter, with the traditional *prediction step* and *update step* combined into a single step. The term $\Phi(k)\hat{\theta}(k-1)$ corresponds to the actuator fault term $\Phi(k)\theta$ in (1a), with θ replaced by its estimate $\hat{\theta}(k-1)$. The extra term $\Upsilon(k)[\hat{\theta}(k) - \hat{\theta}(k-1)]$ is for the purpose of compensating the error caused by $\hat{\theta}(k-1) \neq \theta$. This term is essential for the analysis of the properties of the adaptive Kalman filter in the following sections. It has also been introduced in the deterministic adaptive observer in [12], and its continuous time counterpart in [28].

4 Stability of the adaptive Kalman filter

The main purpose of this section is to prove that the *mathematical expectations* of the state and parameter estimation errors tend exponentially to zero when $k \to \infty$. In other words, the deterministic part of the error dynamics system is exponentially stable.

Another purpose of this section is to prove that the recursively computed matrices $P(k|k), \Upsilon(k), S(k)$ are all bounded, so are the state and parameter estimation gain matrices $K(k)$ and $\Gamma(k)$.

Assumption 1 *The matrices* $A(k), B(k), C(k), \Phi(k)$, $Q(k)$, $R(k)$ *and the input* $u(k)$ *are* upper bounded, $Q(k)$ *is symmetric positive semidefinite, and* $R(k)$ *is symmetric positive-definite with a strictly positive lower bound, for all* $k \geq 0$. □

For controlled systems, boundedness is usually ensured by controllers, in particular the input can be saturated or constrained, like in the case of Model Predictive Control (MPC).

Assumption 2 *The* $[A(k), C(k)]$ *pair is uniformly completely observable, and the* $[A(k), Q^{\frac{1}{2}}(k)]$ *pair is uniformly completely controllable, in the sense of the uniform positive definiteness of the corresponding Gramian matrices. [16, 15].* □

As the lack of observability and/or controllability corresponds to non-minimal state space models, it can be avoided by using minimal state space models.

Assumption 3 *The signals contained in the matrix* $\Phi(k)$ *are persistently exciting in the sense that there exit an integer* $h > 0$ *and a real constant* $\alpha > 0$ *such that, for all integer* $k \geq 0$, *the matrix sequence* $\Omega(k)$ *driven by* $\Phi(k)$ *through the linear system* (5e)-(5f), *satisfies*

$$\sum_{s=0}^{h-1} \Omega^T(k+s)\Sigma^{-1}(k+s)\Omega(k+s) \geq \alpha I_p. \tag{6}$$

□

The generation of $\Omega(k)$ through (5e)-(5f) can be viewed as a linear filter in state-space form, filtering the signals contained in the matrix $\Phi(k)$. When the vector θ has no more than components than output sensors, *i.e.*, $p \leq m$, the matrix $\Omega(k) \in \mathbb{R}^{m \times p}$ has no more columns than rows, then each term in the sum of (6) is typically of full rank. When $p > m$, Assumption 3 means that the signals contained in $\Phi(k)$ must vary sufficiently over the time. In the case of time invariant systems (with constant matrices A, C, K), like the classical persistent excitation conditions for system identification or parameter estimation [18], Assumption 3 implies a sufficient number of frequency components in the signals contained in $\Phi(k)$.

Proposition 1 *The matrices* $P(k|k), \Upsilon(k), \Sigma(k), K(k)$, $\Omega(k)$ *all have a finite upper bound, and* $\Sigma(k)$ *has a strictly positive lower bound.* □

Proof. The proof based on classical results is quite straightforward. The recursive computations (5a)-(5d) for $\Sigma(k), K(k), P(k|k)$ are identical to the corresponding part in the classical Kalman filter, hence like in the classical Kalman filter theory [16, 15], the boundedness of $P(k|k)$ is ensured by the complete uniform observability of the $[A(k), C(k)]$ pair and the complete uniform controllability of the $[A(k), Q^{\frac{1}{2}}(k)]$ pair stated in Assumption 2. As simple corollaries of this result and Assumption 1, the matrices $\Sigma(k)$ and $K(k)$ are also bounded.

Equation (5b) implies $\Sigma(k) > R(k)$. The covariance matrix $R(k)$ is assumed to have a strictly positive lower bound, that is also a lower bound of $\Sigma(k)$.

Again under the complete uniform observability and the complete uniform controllability conditions, the homogenous part of the Kalman filter, corresponding to the homogenous system

$$\zeta(k) = [I_n - K(k)C(k)]A(k)\zeta(k-1), \tag{7}$$

is exponentially stable [16, 15]. This result implies that $\Upsilon(k)$ driven by bounded $\Phi(k)$ through (5e) is bounded. It is then follows from (5f) that $\Omega(k)$ is also bounded. □

Notice that $S(k)$ was missing in Proposition 1. It is the object of the following proposition.

Proposition 2 *The matrix* $S(k)$ *recursively computed through* (5i) *has a finite upper bound and a strictly positive lower bound for all* $k \geq 0$. □

Proof. In order to show the upper and lower bounds of $S(k)$, let us first study another recursively defined matrix sequence

$$M(0) = S^{-1}(0) = \omega^{-1} I_p > 0 \tag{8}$$

$$M(k) = \lambda M(k-1) + \Omega^T(k)\Sigma^{-1}(k)\Omega(k). \tag{9}$$

It will be shown later that $S(k) = M^{-1}(k)$, but for the moment this relationship is not used.

According to Proposition 1, $\Omega(k)$ is upper bounded and $\Sigma(k)$ has a strictly positive lower bound, hence $\Omega^T(k)\Sigma^{-1}(k)\Omega(k)$ is upper bounded. Then $M(k)$ recursively generated from (9) with $\lambda \in (0, 1)$ is also upper bounded. The lower bound of $M(k)$ is investigated in the following.

Repeating the recursion of $M(k)$ in (9) yields

$$M(k) = \lambda^k M(0) + \sum_{j=0}^{k-1} \lambda^j \Omega^T(k-j)\Sigma^{-1}(k-j)\Omega(k-j)$$

For $0 \leq k \leq h$, $M(k) \geq \lambda^h M(0)$.

For $k > h$, let $[k/h]$ denote the largest integer smaller than or equal to k/h, and break the sum $\sum_{j=0}^{k-1}$ into subsums of h terms, then

$$M(k) \geq \sum_{i=1}^{[k/h]} \sum_{j=(i-1)h+0}^{(i-1)h+(h-1)} \lambda^j \Omega^T(k-j)\Sigma^{-1}(k-j)\Omega(k-j)$$
$$\geq \sum_{i=1}^{[k/h]} \lambda^{ih-1} \sum_{j=(i-1)h+0}^{(i-1)h+(h-1)} \Omega^T(k-j)\Sigma^{-1}(k-j)\Omega(k-j)$$
$$\geq \sum_{i=1}^{[k/h]} \lambda^{ih-1}\alpha I_p$$

where the last inequality is based on Assumption 3 (persistent excitation). The geometric sequence sum in this last result can be explicitly computed, so that

$$M(k) \geq \frac{\alpha\lambda^{h-1}(1-\lambda^{h[k/h]})}{1-\lambda^h} I_p \geq \alpha\lambda^{h-1} I_p.$$

Therefore, $M(k)$ has a strictly positive lower bound for either $0 \leq k \leq h$ or $k > h$.

Now take the matrix inverse of both sides of equation (9), and apply the matrix inversion formula

$$(A+V^TBV)^{-1} = A^{-1}-A^{-1}V^T(B^{-1}+VA^{-1}V^T)^{-1}VA^{-1}$$

with $A = \lambda M(k-1)$, $B = \Sigma^{-1}(k)$ and $V = \Omega(k)$, then

$$M^{-1}(k) = \frac{1}{\lambda}M^{-1}(k-1) - \frac{1}{\lambda}M^{-1}(k-1)\Omega^T(k)\big[\lambda\Sigma(k) + \Omega(k)M^{-1}(k-1)\Omega^T(k)\big]^{-1}\Omega(k)M^{-1}(k-1)$$

This recursion in $M^{-1}(k)$ coincides exactly with that of $S(k)$ as formulated in (5i). Moreover, as defined in (8), $M(0) = S^{-1}(0)$, hence $M^{-1}(k) = S(k)$ for *all* $k = 0, 1, 2, \ldots$. It then follows from the already proved upper and lower bounds of $M(k)$ that $S(k)$ has also a finite upper bound and a strictly positive lower bound. □

Define the state and parameter estimation errors

$$\tilde{x}(k|k) \triangleq x(k) - \hat{x}(k|k) \quad (10)$$
$$\tilde{\theta}(k) \triangleq \theta - \hat{\theta}(k) \quad (11)$$

The main result of this section is stated in the following proposition.

Proposition 3 *The mathematical expectations* $\mathrm{E}\tilde{x}(k|k)$ *and* $\mathrm{E}\tilde{\theta}(k)$ *tend to zero exponentially when* $k \to \infty$. □

In other words, the state and parameter estimates of the adaptive Kalman filter converge respectively to the true state $x(x)$ and to the true parameter θ *in mean.* It also means that the deterministic part of the error dynamic system, ignoring the random noise terms, is exponentially stable.

Proof. It is straightforward to compute from (1), (5j) and (5l) that

$$\begin{aligned}\tilde{x}(k|k) &= A(k)\tilde{x}(k-1|k-1) + \Phi(k)\tilde{\theta}(k-1) + w(k) \\ &\quad - K(k)\tilde{y}(k) - \Upsilon(k)[\hat{\theta}(k) - \hat{\theta}(k-1)] \\ &= [I_n - K(k)C(k)][A(k)\tilde{x}(k-1|k-1) + \Phi(k)\tilde{\theta}(k-1)] \\ &\quad + \Upsilon(k)[\tilde{\theta}(k) - \tilde{\theta}(k-1)] \\ &\quad + [I_n - K(k)C(k)]w(k) - K(k)v(k),\end{aligned}$$

and from (5j) and (5k) that

$$\tilde{\theta}(k) = \tilde{\theta}(k-1) - \Gamma(k)\tilde{y}(k) \quad (12)$$
$$= \tilde{\theta}(k-1) - \Gamma(k)C(k)[A(k)\tilde{x}(k-1|k-1) + \Phi(k)\tilde{\theta}(k-1)] - \Gamma(k)C(k)w(k) - \Gamma(k)v(k). \quad (13)$$

Like in [28], define

$$\xi(k) \triangleq \tilde{x}(k|k) - \Upsilon(k)\tilde{\theta}(k). \quad (14)$$

Simple substitutions lead to

$$\begin{aligned}\xi(k) = {} & [I_n - K(k)C(k)][A(k)\tilde{x}(k-1|k-1) + \Phi(k)\tilde{\theta}(k-1)] \\ & + \Upsilon(k)[\tilde{\theta}(k) - \tilde{\theta}(k-1)] \\ & + [I_n - K(k)C(k)]w(k) - K(k)v(k) \\ & - \Upsilon(k)\tilde{\theta}(k).\end{aligned}$$

In this last result, according to (14), replace $\tilde{x}(k-1|k-1)$ with

$$\tilde{x}(k-1|k-1) = \xi(k-1) + \Upsilon(k-1)\tilde{\theta}(k-1), \quad (15)$$

then

$$\begin{aligned}\xi(k) = {} & [I_n - K(k)C(k)]A(k)\xi(k-1) \\ & + \big\{[I_n - K(k)C(k)]A(k)\Upsilon(k-1) \\ & + [I_n - K(k)C(k)]\Phi(k) - \Upsilon(k)\big\}\tilde{\theta}(k-1)] \\ & + [I_n - K(k)C(k)]w(k) - K(k)v(k).\end{aligned}$$

The content of the curly braces $\{\cdots\}$ is zero, because $\Upsilon(k)$ satisfies (5e). Then

$$\begin{aligned}\xi(k) = {} & [I_n - K(k)C(k)]A(k)\xi(k-1) \\ & + [I_n - K(k)C(k)]w(k) - K(k)v(k).\end{aligned} \quad (16)$$

The noises $w(k)$ and $v(k)$ are assumed to have zero mean values (centered noises), then

$$\mathrm{E}\xi(k) = [I_n - K(k)C(k)]A(k)\mathrm{E}\xi(k-1).$$

Like (7), this recurrent equation is exponentially stable, hence $\mathrm{E}\xi(k) \to 0$ for any initial value $\mathrm{E}\xi(0)$. In fact, starting from

$$\mathrm{E}\xi(0) = \mathrm{E}\tilde{x}(0) - \Upsilon(0)\mathrm{E}\tilde{\theta}(0) = 0,$$

it is recursively shown that $\mathrm{E}\xi(k) = 0$ *for all* $k \geq 0$.
In (13) replace $\tilde{x}(k-1|k-1)$ with (15), then

$$\begin{aligned}\tilde{\theta}(k) &= \big[I_p - \Gamma(k)C(k)A(k)\Upsilon(k-1) \\ &\quad - \Gamma(k)C(k)\Phi(k)\big]\tilde{\theta}(k-1) - \Gamma(k)C(k)A(k)\xi(k-1) \\ &\quad - \Gamma(k)C(k)w(k) - \Gamma(k)v(k) \\ &= [I_p - \Gamma(k)\Omega(k)]\tilde{\theta}(k-1) - \Gamma(k)e(k) \end{aligned} \tag{17}$$

where $\Omega(k)$ is as defined in (5f), and

$$e(k) \triangleq C(k)A(k)\xi(k-1) + C(k)w(k) + v(k). \tag{18}$$

It was already shown $\mathrm{E}\xi(k) = 0$ for all $k \geq 0$, therefore $\mathrm{E}e(k) = 0$.
Take the mathematical expectation at both sides of (17) and denote

$$\bar{\tilde{\theta}}(k) \triangleq \mathrm{E}\tilde{\theta}(k), \tag{19}$$

then

$$\bar{\tilde{\theta}}(k) = [I_p - \Gamma(k)\Omega(k)]\,\bar{\tilde{\theta}}(k-1). \tag{20}$$

Before analyzing the convergence of $\bar{\tilde{\theta}}(k)$ governed by (20), let us combine the two equations (5h) and (5i) into

$$S(k) = \frac{1}{\lambda}[I_p - \Gamma(k)\Omega(k)]S(k-1). \tag{21}$$

Accordingly, $M(k) = S^{-1}(k)$ satisfies [2]

$$M(k) = \lambda M(k-1)[I_p - \Gamma(k)\Omega(k)]^{-1}. \tag{22}$$

Let us define the Lyapunov function candidate

$$V(k) \triangleq \left(\bar{\tilde{\theta}}(k)\right)^T M(k)\bar{\tilde{\theta}}(k), \tag{23}$$

it then follows from (20) and (22) that

$$\begin{aligned} V(k) &= \left(\bar{\tilde{\theta}}(k-1)\right)^T [I_p - \Gamma(k)\Omega(k)]^T \lambda M(k-1)\bar{\tilde{\theta}}(k-1) \\ &= \lambda\left(\bar{\tilde{\theta}}(k-1)\right)^T M(k-1)\bar{\tilde{\theta}}(k-1) \\ &\quad - \lambda\left(\bar{\tilde{\theta}}(k-1)\right)^T \Omega^T(k)\Gamma^T(k)M(k-1)\bar{\tilde{\theta}}(k-1). \end{aligned} \tag{24}$$

[2] According to Proposition 2, $S(k)$ is positive definite for all $k \geq 0$, hence (21) implies that the matrix $[I_p - \Gamma(k)\Omega(k)]$ is invertible.

By recalling (5h), (5g) and $M(k-1) = S^{-1}(k-1)$, it yields

$$\begin{aligned} \Xi(k) &\triangleq \Omega^T(k)\Gamma^T(k)M(k-1) & (25)\\ &= \Omega^T(k)\Lambda(k)\Omega(k) & (26)\\ &= \Omega^T(k)\big[\lambda\Sigma(k) + \Omega(k)S(k-1)\Omega^T(k)\big]^{-1}\Omega(k), & (27)\end{aligned}$$

which is a symmetric positive semidefinite matrix. Then (24) becomes

$$\begin{aligned} V(k) &= \lambda V(k-1) - \lambda\left(\bar{\tilde{\theta}}(k-1)\right)^T \Xi(k)\bar{\tilde{\theta}}(k-1) \\ &\leq \lambda V(k-1), \end{aligned} \tag{28}$$

implying that $V(k) = \left(\bar{\tilde{\theta}}(k)\right)^T M(k)\bar{\tilde{\theta}}(k)$ converges to zero exponentially. It is already shown that $M(k)$ is lower and upper bounded, with a strictly positive lower bound, $\mathrm{E}\tilde{\theta}(k) = \bar{\tilde{\theta}}(k)$ then converges to zero exponentially.
Finally, it follows from (14) that

$$\mathrm{E}\tilde{x}(k|k) = \mathrm{E}\xi(k) + \Upsilon(k)\mathrm{E}\tilde{\theta}(k) = \Upsilon(k)\mathrm{E}\tilde{\theta}(k), \tag{29}$$

it is then concluded that the mathematical expectations $\mathrm{E}\tilde{x}(k|k)$ and $\mathrm{E}\tilde{\theta}(k)$ tend to zero exponentially when $k \to \infty$. □

In addition to the already established fact that $\mathrm{E}e(k) = 0$, the following lemma characterizes more completely the properties of the error sequence $e(k)$ defined in (18). This lemma will be useful in Section 6.

Lemma 1 *The sequence $e(k) \in \mathbb{R}^m$ defined in* (18) *is identical to the innovation sequence $\varepsilon(k)$ of the* standard *Kalman filter applied to the* fault-free *(corresponding to $\theta = 0$) system* (1)*, which is a white Gaussian noise with its covariance matrix $\Sigma(k)$ as computed in* (5b). □

The proof of this lemma is *non trivial.* It is presented in Appendix A in order not to shade the main results of this paper.

5 Minimum covariance of combined estimation errors

The following result is a generalization of the minimum variance property of the classical Kalman filter.

Proposition 4 *In the adaptive Kalman filter* (5)*, relax the Kalman gain $K(k)$ computed through the recurrent equations* (5a)-(5d) *to* any *matrix sequence $L(k) \in \mathbb{R}^{n\times m}$. Consider the combined state and parameter estimation error $\xi(k) = \tilde{x}(k|k) - \Upsilon(k)\tilde{\theta}(k)$. Its covariance matrix depending on the gain sequence $L(k)$ and denoted by* $\mathrm{cov}[\xi(k)|L]$ *reaches its minimum when $L(k) = K(k)$, in the sense of the positive definiteness of the difference matrix:*

$$\mathrm{cov}[\xi(k)|L] - \mathrm{cov}[\xi(k)|K] \geq 0 \tag{30}$$

for any $L(k) \in \mathbb{R}^{n\times m}$. $\square$

This result means that $v^T\mathrm{cov}[\xi(k)|L]v \geq v^T\mathrm{cov}[\xi(k)|K]v$ for *any* vector $v \in \mathbb{R}^n$. In fact, the term $\Upsilon(k)\tilde{\theta}(k)$ is the part of the state estimation error $\tilde{x}(k|k)$ due to the parameter estimation error $\tilde{\theta}(k)$ (this part would be zero if the parameter estimate $\hat{\theta}(k)$ was replaced by the true parameter value θ, and then the adaptive Kalman filter (5) would be reduced to the standard Kalman filter). Therefore, the meaning of this proposition is that the remaining part of the state estimation error reaches its minimum variance if the Kalman gain $K(k)$ is used.

Proof.

Following (16) where $K(k)$ is replaced by $L(k)$, and noticing that $\xi(k-1), w(k)$ and $v(k)$ are pairwise independent, compute the covariance matrix of $\xi(k)$:

$$\begin{aligned}\mathrm{cov}[\xi(k)|L] &= \mathrm{E}[\xi(k)\xi^T(k)]\\ &= [I_n - L(k)C(k)]A(k)\mathrm{cov}[\xi(k-1)|L]\\ &\quad \cdot A^T(k)[I_n - L(k)C(k)]^T\\ &\quad + [I_n - L(k)C(k)]Q(k)[I_n - L(k)C(k)]^T\\ &\quad + L(k)R(k)L^T(k) \end{aligned} \tag{31}$$

For shorter notations, let us denote

$$\Pi(k) \triangleq A(k)\mathrm{cov}[\xi(k-1)|L]A^T(k) + Q(k), \tag{32}$$

which is independent of $L(k)$ (of course, $\Pi(k)$ depends on $L(k-1)$). Then

$$\begin{aligned}\mathrm{cov}[\xi(k)|L] = [I_n - L(k)C(k)]\Pi(k)[I_n - L(k)C(k)]^T\\ + L(k)R(k)L^T(k).\end{aligned} \tag{33}$$

Define also

$$H(k) \triangleq C(k)\Pi(k)C^T(k) + R(k). \tag{34}$$

Because $C(k)\Pi(k)C^T(k) \geq 0$ and $R(k) > 0$ ($R(k)$ is a positive definite matrix, see Assumption 1), $H(k)$ is also *positive definite*, and thus invertible.

Rearrange (33) as

$$\begin{aligned}\mathrm{cov}[\xi(k)|L] &= [L(k) - \Pi(k)C^T(k)H^{-1}(k)]H(k)\\ &\quad \cdot [L(k) - \Pi(k)C^T(k)H^{-1}(k)]^T + \Pi(k)\\ &\quad - \Pi(k)C^T(k)H^{-1}(k)C(k)\Pi(k).\end{aligned} \tag{35}$$

The equivalence between (33) and (35) can be shown by first developing (35) and then by incorporating (34).

The matrix $H(k)$ is *positive definite*, hence the first term in (35) (a symmetric matrix product) is positive semidefinite, that is,

$$\begin{aligned}&[L(k) - \Pi(k)C^T(k)H^{-1}(k)]H(k)\\ &\quad \cdot [L(k) - \Pi(k)C^T(k)H^{-1}(k)]^T \geq 0.\end{aligned}$$

When $L(k)$ is chosen such that this inequality becomes equality, *i.e.*, the first term of (35) is zero, $\mathrm{cov}[\xi(k)]$ reaches its minimum, since the other terms in (35) are independent of $L(k)$. This optimal choice of $L(k)$, denoted by $L_*(k)$, is

$$L_*(k) \triangleq \Pi(k)C^T(k)H^{-1}(k). \tag{36}$$

It remains to show that $L_*(k)$ is identical to the Kalman gain $K(k)$ in order to complete the proof.

Rewrite (33) while incorporating (34):

$$\begin{aligned}\mathrm{cov}[\xi(k)|L] = \Pi(k) - L(k)C(k)\Pi(k) - \Pi(k)C^T(k)L^T(k)\\ + L(k)H(k)L^T(k).\end{aligned} \tag{37}$$

In the particular case $L(k) = L_*(k) = \Pi(k)C^T(k)H^{-1}(k)$,

$$\begin{aligned}\mathrm{cov}[\xi(k)|L_*] &= \Pi(k) - 2\Pi(k)C^T(k)H^{-1}(k)C(k)\Pi(k)\\ &\quad + \Pi(k)C^T(k)H^{-1}(k)H(k)[\Pi(k)C^T(k)H^{-1}(k)]^T.\\ &= \Pi(k) - \Pi(k)C^T(k)H^{-1}(k)C(k)\Pi(k) \qquad (38)\\ &= [I_n - L_*(k)C(k)]\Pi(k) \qquad (39)\end{aligned}$$

Assemble (32), (34), (36) and (39) together, for $L = L_*$,

$$\Pi(k) = A(k)\mathrm{cov}[\xi(k-1)|L_*]A^T(k) + Q(k) \tag{40a}$$

$$H(k) = C(k)\Pi(k)C^T(k) + R(k) \tag{40b}$$

$$L_*(k) = \Pi(k)C^T(k)H^{-1}(k) \tag{40c}$$

$$\mathrm{cov}[\xi(k)|L_*] = [I_n - L_*(k)C(k)]\Pi(k). \tag{40d}$$

These equations allow recursive computation of $\mathrm{cov}[\xi(k)|L_*]$ and the other involved matrices. It turns out that these recursive computations are *exactly the same* as those in (5a)-(5d), with $\Pi(k), H(k), L_*(k)$ and $\mathrm{cov}[\xi(k)|L_*]$ corresponding respectively to $P(k|k-1), \Sigma(k), K(k)$ and $P(k|k)$.

It remains to show that these two recursive computations have the same initial condition, *i.e.* $\mathrm{cov}[\xi(0)|L_*] = P(0|0)$, in order to conclude $L_*(k) = K(k)$ for all $k \geq 0$.

Because $\Upsilon(0) = 0$ as specified in (4a), and according to the definition of $\xi(k)$ in (14),

$$\xi(0) = \tilde{x}(0|0) - \Upsilon(0)\tilde{\theta}(0) = \tilde{x}(0|0), \tag{41}$$

hence

$$\mathrm{cov}[\xi(0)|L_*] = \mathrm{cov}[\tilde{x}(0|0)] = P(0|0). \tag{42}$$

Therefore, $L_*(k) = K(k)$ for all $k \geq 0$.

It is then established that the covariance matrix $\mathrm{cov}[\xi(k)|L]$ reaches its minimum when $L(k) = K(k)$. $\square$

6 Equivalence to recursive least squares parameter estimator

The following proposition states that the parameter estimation within the adaptive Kalman filter is equivalent to the classical recursive least squares (RLS) estimation formulated for a fictive regression problem.

Proposition 5 *Consider the linear regression*

$$z(k) = \Omega(k)\theta + \varepsilon(k) \tag{43}$$

where

- *the matrix of regressors* $\Omega(k) \in \mathbb{R}^{m\times p}$ *is as defined in* (5f)*,*
- *the regression parameter vector* $\theta \in \mathbb{R}^p$ *is equal to the vector* θ *appearing in the term* $\Phi(k)\theta$ *of system* (1)*,*
- *the error term* $\varepsilon(k) \in \mathbb{R}^m$ *is equal to the* innovation sequence *of the* standard *Kalman filter applied to the* fault-free *(*$\theta = 0$*) system* (1)*, which is a white Gaussian noise with its covariance matrix* $\Sigma(k)$ *as computed in* (5b)*.*

Then the RLS estimate $\hat\theta_{RLS}(k)$ *for the linear regression* (43) *is identical to the parameter estimate* $\hat\theta(k)$ *computed in* (5k)*, i.e.,* $\hat\theta_{RLS}(k) = \hat\theta(k)$ *for all* $k \geq 0$*, provided the two algorithms are appropriately initialized and use the same forgetting factor* λ*.* □

Proof. The classical RLS estimator (see, *e.g.*, [18]) applied to the linear regression (43) is

$$\Lambda(k) = [\lambda\Sigma(k) + \Omega(k)S(k-1)\Omega^T(k)]^{-1} \tag{44a}$$
$$\Gamma(k) = S(k-1)\Omega^T(k)\Lambda(k) \tag{44b}$$
$$S(k) = \frac{1}{\lambda}S(k-1) - \frac{1}{\lambda}S(k-1)\Omega^T(k)\Lambda(k)\Omega(k)S(k-1) \tag{44c}$$
$$\hat\theta_{\mathrm{RLS}}(k) = \hat\theta_{\mathrm{RLS}}(k-1) + \Gamma(k)[z(k) - \Omega(k)\hat\theta_{\mathrm{RLS}}(k-1)] \tag{44d}$$

with the initial values $\hat\theta_{\mathrm{RLS}}(0) = \theta_0$ and $S(0) = \omega I_p$, chosen as those appearing in (4a) and (4b).

Because $\Omega(k)$, $\Sigma(k)$ and λ are the same as those appearing in (5), so are $\Lambda(k), S(k)$ and $\Gamma(k)$ computed with identical formulas (the initial value $S(0) = \omega I_p$ is also assumed identical in the two algorithms).

Define

$$\tilde\theta_{\mathrm{RLS}}(k) \triangleq \theta - \hat\theta_{\mathrm{RLS}}(k), \tag{45}$$

then (44d) leads to

$$\tilde\theta_{\mathrm{RLS}}(k) = \tilde\theta_{\mathrm{RLS}}(k-1) - \Gamma(k)[z(k) - \Omega(k)\hat\theta_{\mathrm{RLS}}(k-1)]$$

Substitute $z(k)$ with (43), then

$$\begin{aligned}\tilde\theta_{\mathrm{RLS}}(k) &= \tilde\theta_{\mathrm{RLS}}(k-1) - \Gamma(k)[\Omega(k)\tilde\theta_{\mathrm{RLS}}(k-1) + \varepsilon(k)] \\ &= [I_p - \Gamma(k)\Omega(k)]\tilde\theta_{\mathrm{RLS}}(k-1) - \Gamma(k)\varepsilon(k).\end{aligned} \tag{46}$$

On the other hand, for the adaptive Kalman filter, the parameter estimate error $\tilde\theta(k)$ was defined in (11), and it was shown that $\tilde\theta(k)$ satisfies the recurrent equation (17), which is similar to (46) satisfied by $\tilde\theta_{\mathrm{RLS}}(k)$. According to Lemma 1 formulated at the end of Section 4, $e(k) = \varepsilon(k)$ for all $k \geq 0$, hence indeed the two recurrent equations (17) and (46) are identical.

Moreover, for the initialization of the two algorithms it has been chosen that $\hat\theta(0) = \hat\theta_{\mathrm{RLS}}(0) = \theta_0$, hence $\tilde\theta(0) = \tilde\theta_{\mathrm{RLS}}(0)$. Therefore, $\tilde\theta(k) = \tilde\theta_{\mathrm{RLS}}(k)$ for all $k \geq 0$. As $\hat\theta(k) = \theta - \tilde\theta(k)$ and $\hat\theta_{\mathrm{RLS}}(k) = \theta - \tilde\theta_{\mathrm{RLS}}(k)$, it is then concluded that $\hat\theta(k) = \hat\theta_{\mathrm{RLS}}(k)$ for all $k \geq 0$. □

This result allows to well understand the role played by the forgetting factor λ. Like in the classical RLS algorithm, it controls how fast past observations are forgotten. As illustrated by the second numerical example in the following section, the smaller λ is, the faster past observations are forgotten, and the faster the transient behavior of the algorithm is, but the more sensitive the result is to noises.

7 Numerical examples

Randomly generated examples are first presented to illustrate the statistical properties of the proposed adaptive Kalman filter, before a more concrete example.

7.1 Randomly generated examples

Consider a piecewise constant system randomly switching within 4 third order ($n = 3$) state-space models with one input ($l = 1$) and 2 outputs ($m = 2$). Each of the 4 state-space models is randomly generated with the Matlab code (requiring the System Identification Toolbox):

```
zreal = rand(1,1)*1.2-0.6; pmodul = rand(1,1)*0.1+0.4;
pphase = rand(1,1)*2*pi;    preal = rand(1,1)-0.5;
ssk = idss(zpk(zreal,[pmodul.*exp(1i*pphase) ...
  pmodul.*exp(-1i*pphase) preal],rand(1,1)+0.5,1));
```

In the simulation for $k = 0, 1, 2, \ldots, 1000$, the actual model at each instant k randomly switches among the 4 randomly generated state-space models, which are kept unchanged. The random switching sequence among the 4 state-space models is plotted in Figure 1. The input $u(k)$ is randomly generated with a Gaussian distribution and the standard deviation equal to 2. The noise covariance matrices are chosen as $Q(k) = 0.1I_3$ and $R(k) = 0.05I_2$ for $k \geq 0$. The matrix $\Phi(k)$ is as in (2) so that θ represents actuator gain loss. During the numerical simulation running from $k = 0$ to $k = 1000$, a gain loss of 50% at the time instant $k = 500$ is simulated, corresponding to a jump of θ (a scalar parameter) from 0 to 0.5. The

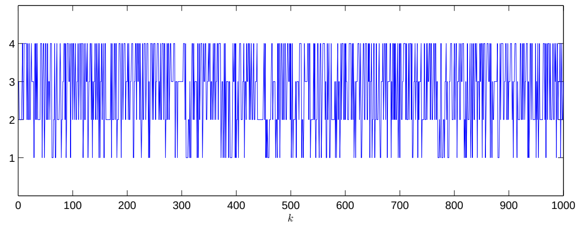

Fig. 1. Random switching index among 4 state-space models.

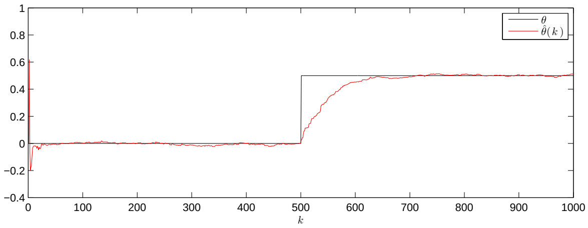

Fig. 2. The simulated "true" parameter θ and parameter estimate $\hat{\theta}(k)$ by the adaptive Kalman filter.

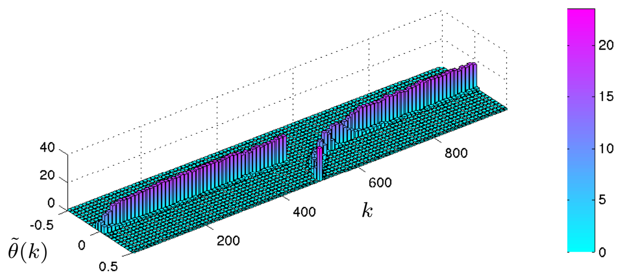

Fig. 3. Histogram per instant k of the parameter estimation error of the adaptive Kalman filter.

result of parameter estimation by the adaptive Kalman filter is presented in Fig. 2.

The above results are based on a single numerical simulation trial. In order to statistically evaluate the performance of the proposed method, 1000 simulated trials are performed, each corresponding to a different set of 4 state-space models and to a different random realization of the input and noises. At each time instant k, the histogram of the parameter estimation error based on the 1000 simulated trials is generated, and all the histograms are depicted as a 3D illustration in Fig, 3. The histograms are normalized so that they are similar to probability density functions.

7.2 Lateral dynamics of a remotely piloted aircraft

Consider a remotely piloted aircraft as described in [6, page 188]. Its linearized lateral dynamics discretized with the sampling period of 0.1s is described by a state-space equation with

$$x(k) = \begin{bmatrix} \text{side slip} \\ \text{roll rate} \\ \text{yaw rate} \\ \text{bank angle} \\ \text{yaw angle} \end{bmatrix}, \; u(k) = \begin{bmatrix} \text{rudder} \\ \text{aileron} \end{bmatrix},$$

$$A = \begin{bmatrix} 0.9157 & -0.0369 & -3.0853 & -0.9486 & 0 \\ -0.0017 & 0.4264 & 0.2500 & 0.0020 & 0 \\ 0.0342 & -0.0005 & 0.8816 & -0.0172 & 0 \\ -0.0002 & 0.0673 & 0.0144 & 1.0001 & 0 \\ 0.0018 & -0.0000 & 0.0950 & -0.0006 & 1.0000 \end{bmatrix},$$

$$B = \begin{bmatrix} 1.5686 & 0 \\ -0.1751 & 1.2208 \\ 0.0204 & 0 \\ 0 & -0.1432 \\ -0.0474 & 0 \end{bmatrix}, \; C = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

The two actuators (rudder and aileron) are subject to failures leading to gain losses, respectively represented by the two components θ_1 and θ_2 of the parameter vector θ. Accordingly the matrix $\Phi(k)$ is expressed as in (2).

Numerical simulations are made for $k = 1, 2, \ldots, 1000$. At the beginning $\theta_1 = \theta_2 = 0$. At $k = 300$, θ_1 changes from 0 to 0.2, and at $k = 600$, θ_2 changes from 0 to 0.1. The other simulation parameters are: the state and output noise covariance matrices $Q = 0.01I_5$, $R = 0.01I_3$, the initial state $x_0 = 0$, the covariance of the initial state $P_0 = I_5$, the initial matrix $S(0) = I_2$, the initial parameter estimate $\hat{\theta}(0) = 0$, and the two inputs are randomly generated. In order to illustrate the effect of the forgetting factor λ, three simulations are made with $\lambda = 0.9, 0.97$ and 0.99. The results of the adaptive Kalman filter corresponding to these different values of λ are shown in Figure 4, where the solid lines represent the parameter estimates $\hat{\theta}_1(k)$ and $\hat{\theta}_2(k)$, and the dashed lines indicate the true actuator gain losses of the simulator. After each parameter change, the parameter estimates converge after a transient depending on the forgetting factor λ: a smaller λ leads to a faster transient with a higher sensitivity to noises, whereas a larger λ results in a slower transient with smoother estimates.

8 Conclusion

Unlike classical adaptive Kalman filters, which have been designed for state estimation in case of uncertainties about noise covariances, the adaptive Kalman filter proposed in this paper is for the purpose actuator fault diagnosis, through joint state-parameter estimation. It is applicable to LTV/LPV systems. The stability and

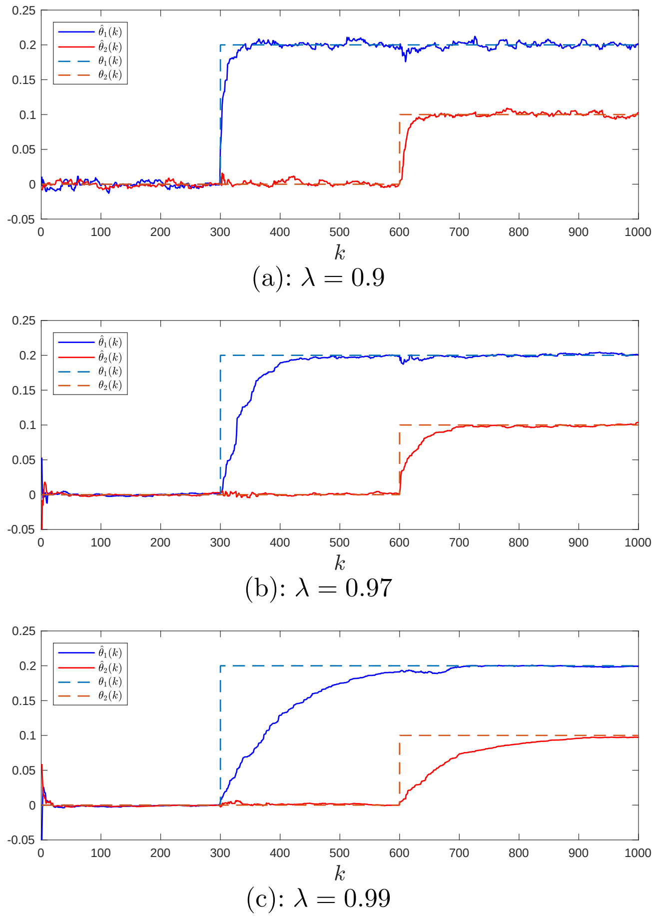

Fig. 4. Simulated remotely piloted aircraft actuator fault diagnosis with different forgetting factors, (a): $\lambda = 0.9$, (b): $\lambda = 0.97$, (c): $\lambda = 0.99$. Solid lines: estimates $\hat{\theta}_1(k)$ and $\hat{\theta}_2(k)$ of actuator gain losses, dashed lines: true actuator gain losses of the simulator.

minimum variance properties of the adaptive Kalman filter have been rigorously analyzed. Through LTV/LPV reformulation and approximations, this method for actuator fault diagnosis is also applicable to a large class of nonlinear systems.

The basic structure of this algorithm is not new, because it has been used in some variant forms in deterministic frameworks as adaptive observers. However, to the author's knowledge, this paper presents the first version in a stochastic framework, with rigorously established statistical and stability properties. In particular, it is shown that the proposed adaptive Kalman filter provides a recursive estimation of actuator fault parameters equivalent to the recursive least squares estimator formulated for a fictive regression problem.

By restricting the parameter vector θ to a finite set of possible values, it would be possible to address the joint state-parameter estimation problem as a hybrid system filtering problem. Such a multi-mode formulation would have the advantage of better dealing with frequent parameter changes. As actuator faults are typically rare events, the diagnosis method proposed in this paper is a better trade-off between algorithm simplicity and efficiency.

A Proof of Lemma 1

Warning on notations. In this appendix, the notations like $x(k), \hat{x}(k|k), \hat{x}(k|k-1)$ etc., unless otherwise specified, are about the *fault-free* system and its *standard* Kalman filter, to be distinguished from those in the main body of this paper about the system subject to actuator faults and its *adaptive* Kalman filter.

The *standard* Kalman filter (as known in classical textbooks, different from the *adaptive* Kalman filter presented in this paper) is applied to the *fault-free* system, characterized by the state space model (1) with the term $\Phi(k)\theta$ being omitted. More clearly, the state equation of the *fault-free* system is

$$x(k) = A(k)x(k-1) + B(k)u(k) + w(k). \tag{A.1}$$

This Kalman filter is based on the same gain $K(k)$ and on the related matrix sequences as computed in (5a)-(5d) within the adaptive Kalman filter (5).

State estimation is usually expressed in two steps, namely *prediction* and *update*, as

$$\hat{x}(k|k-1) = A(k)\hat{x}(k-1|k-1) + B(k)u(k) \tag{A.2a}$$
$$\hat{x}(k|k) = \hat{x}(k|k-1) + K(k)\varepsilon(k) \tag{A.2b}$$

where the innovation $\varepsilon(k)$ is defined as

$$\varepsilon(k) \triangleq y(k) - C(k)\hat{x}(k|k-1). \tag{A.3}$$

It is known from the classical Kalman filter theory that this innovation sequence is a white Gaussian noise with its covariance matrix equal to $\Sigma(k)$ as computed in (5b), if the Kalman gain $K(k)$ and the related matrix sequences are computed as in (5a)-(5d) from the system matrix $A(k)$ and the noise covariance matrices $Q(k)$ and $R(k)$.

Eliminating $\hat{x}(k|k)$ from the recursions (A.2) yields

$$\begin{aligned}\hat{x}(k|k-1) = A(k)\hat{x}(k-1|k-2) + A(k)K(k-1)\varepsilon(k-1)\\ + B(k)u(k)\end{aligned} \tag{A.4}$$

Subtract this equation from the corresponding sides of equation (A.1), then

$$\begin{aligned}\tilde{x}(k|k-1) = A(k)\tilde{x}(k-1|k-2) - A(k)K(k-1)\varepsilon(k-1)\\ + w(k),\end{aligned} \tag{A.5}$$

where

$$\tilde{x}(k|k-1) \triangleq x(k) - \hat{x}(k|k-1). \tag{A.6}$$

Substitute $y(k)$ with (1b) in (A.3), yielding

$$\varepsilon(k) = C(k)\tilde{x}(k|k-1) + v(k), \tag{A.7}$$

and insert this result in (A.5), then

$$\begin{aligned}\tilde{x}(k|k-1) &= A(k)[I_n - K(k-1)K(k-1)C(k-1)] \\ &\cdot \tilde{x}(k-1|k-2) - A(k)K(k-1)v(k-1) + w(k),\end{aligned} \tag{A.8}$$

The two equations (A.7) and (A.8) charaterize the standard Kalman innovation $\varepsilon(k)$ driven by the two noises $w(k)$ and $v(k)$. This characterization of $\varepsilon(k)$ is still far from the definition of $e(k)$ in (18).

Plug (A.8) into (A.7), then

$$\begin{aligned}\varepsilon(k) &= C(k)\big\{A(k)[I_n - K(k-1)K(k-1)C(k-1)] \\ &\quad \cdot \tilde{x}(k-1|k-2) - A(k)K(k-1)v(k-1) \\ &\quad + w(k)\big\} + v(k) \\ &= C(k)A(k)\big\{[I_n - K(k-1)K(k-1)C(k-1)] \\ &\quad \cdot \tilde{x}(k-1|k-2) - K(k-1)v(k-1)\big\} \\ &\quad + C(k)w(k) + v(k).\end{aligned} \tag{A.9}$$

Denote

$$\tilde{x}(k|k) \triangleq x(k) - \hat{x}(k|k). \tag{A.10}$$

Subtract $x(k)$ from both sides of (A.2b), then

$$-\tilde{x}(k|k) = -\tilde{x}(k|k-1) + K(k)\varepsilon(k). \tag{A.11}$$

Inverse the signs of the two sides of this equality, and replace $\varepsilon(k)$ with (A.7), then

$$\tilde{x}(k|k) = [I_n - K(k)C(k)]\tilde{x}(k|k-1) - K(k)v(k). \tag{A.12}$$

Replacing k by $k-1$ yields

$$\begin{aligned}\tilde{x}(k-1|k-1) &= [I_n - K(k-1)C(k-1)]\tilde{x}(k-1|k-2) \\ &\quad - K(k-1)v(k-1).\end{aligned} \tag{A.13}$$

This result allows to continue (A.9) as

$$\varepsilon(k) = C(k)A(k)\tilde{x}(k-1|k-1) + C(k)w(k) + v(k). \tag{A.14}$$

This expression of $\varepsilon(k)$ is similar to that of $e(k)$ in (18). It remains to show that $\tilde{x}(k|k) = \xi(k)$ for all $k \geq 0$ in order to prove that $\varepsilon(k)$ and $e(k)$ are identical.

Subtract (A.2a) from the respective sides of (A.1), then

$$\tilde{x}(k|k-1) = A(k)\tilde{x}(k-1|k-1) + w(k). \tag{A.15}$$

Inserting this result into (A.12) yields

$$\begin{aligned}\tilde{x}(k|k) &= [I_n - K(k)C(k)]A(k)\tilde{x}(k-1|k-1) \\ &\quad + [I_n - K(k)C(k)]w(k) - K(k)v(k).\end{aligned} \tag{A.16}$$

This recurrent equation satisfied by $\tilde{x}(k|k)$ is exactly the same as (16) satisfied by $\xi(k)$. It remains to check if their initial values $\xi(0)$ and $\tilde{x}(0|0)$ are identical.

Because $\Upsilon(0) = 0$ as specified in (4a), and according to the definition of $\xi(k)$ in (14),

$$\xi(0) = \tilde{x}(0|0) - \Upsilon(0)\tilde{\theta}(0) = \tilde{x}(0|0). \tag{A.17}$$

Recall that, in this appendix, the notations like $x(k), \hat{x}(k|k)$ and $\hat{x}(k|k-1)$ refer to the *fault-free* system and the *standard* Kalman filter applied to it, so do $\tilde{x}(k|k)$ and $\tilde{x}(k|k-1)$, except those in (A.17), which are indeed about the system subject to actuator faults formulated in (1) and its *adaptive* Kalman filter (5), because (A.17) is derived from (14), which was introduced in the analysis of the adaptive Kalman filter. By assuming the same initial state $x(0)$ for both system (1) and the fault-free system, and by choosing the same initial state estimate $\hat{x}(0|0)$ for both adaptive and standard Kalman filters, the two initial estimation errors then have the same value $\tilde{x}(0|0) = x(0) - \hat{x}(0|0)$. Therefore, the two sequences $\xi(k)$ and $\tilde{x}(k|k)$ generated by the same recurrent equation (16) and (A.16) with the same initial value are indeed equal for all $k \geq 0$. This result completes the proof that $\varepsilon(k) = e(k)$ for all $k \geq 0$, by recalling (18) and (A.14).

In practice, only the input and output of system (1) are available, whereas the fault-free system, formulated for the purpose of analysis, is fictive. There is no need to really implement the standard Kalman filter for the fault-free system. The purpose of this lemma is to show that the error sequence $e(k)$ appearing in equation (17) is a white Gaussian noise of covariance matrix $\Sigma(k)$, as it is equivalent to the innovation sequence of the standard Kalman filter.

□

References

[1] Marouane Alma and Mohamed Darouach. Adaptive observers design for a class of linear descriptor systems. *Automatica*, 50(2):578–583, 2014.

[2] M. Basseville and I. Nikiforov. *Detection of Abrupt Changes - Theory and Application.* Prentice Hall, Englewood Cliffs, 1993.

[3] D. Berdjag, C. Christophe, V. Cocquempot, and B. Jiang. Nonlinear model decomposition for robust fault detection and isolation using algebraic tools. *Int. Jal Innovative Computing, Information and Control*, 2(6):1337–1354, 2006.

[4] G. Besançon, J. De Leon-Morales, and O. Huerta-Guevara. On adaptive observers for state affine systems. *International Journal of Control*, 79(6):581–591, 2006.

[5] Steven D. Brown and Sarah C. Rutan. Adaptive Kalman filtering. *Journal of Research of the National Bureau of Standards*, 90(6):403–407, 1985.

[6] J. Chen and R.J. Patton. *Robust Model-Based Fault Diagnosis for Dynamic Systems.* Kluwer, Boston, 1999.

[7] C. De Persis and A. Isidori. A geometric approach to nonlinear fault detection and isolation. *IEEE Trans. Automatic Control*, 46(6):853–865, 2001.

[8] S.X. Ding. *Model-based Fault Diagnosis Techniques: Design Schemes, Algorithms, and Tools.* Springer, Berlin, 2008.

[9] M. Farza, M. M'Saada, T. Maatouga, and M. Kamounb. Adaptive observers for nonlinearly parameterized class of nonlinear systems. *Automatica*, 45(10):2292–2299, 2009.

[10] Mondher Farza, Ibtissem Bouraoui, Tomas Menard, Ridha Ben Abdennour, and Mohammed M'Saad. Adaptive observers for a class of uniformly observable systems with nonlinear parametrization and sampled outputs. *Automatica*, 50(11):2951–2960, 2014.

[11] J.J. Gertler. *Fault Detection and Diagnosis in Engineering Systems.* Marcel Dekker, New York, 1998.

[12] Arnaud Guyader and Qinghua Zhang. Adaptive observer for discrete time linear time varying systems. In *13th IFAC/IFORS Symposium on System Identification (SYSID)*, Rotterdam, 2003.

[13] I. Hwang, S. Kim, Y. Kim, and C.E. Seah. A survey of fault detection, isolation, and reconfiguration methods. *IEEE Trans. Control Systems Technology*, 18(3):636–653, 2010.

[14] R. Isermann. *Fault Diagnosis Systems: An Introduction From Fault Detection To Fault Tolerance.* Springer, Berlin, 2005.

[15] A. H. Jazwinski. *Stochastic Processes and Filtering Theory*, volume 64 of *Mathematics in Science and Engineering.* Academic Press, New York, 1970.

[16] R. E. Kalman. New methods in Wiener filtering theory. In J. L. Bogdanoff and F. Kozin, editors, *Proceedings of the First Symposium on Engineering Applications of Random Function Theory and Probability*, New York, 1963. John Wiley & Sons.

[17] Yusheng Liu. Robust adaptive observer for nonlinear systems with unmodeled dynamics. *Automatica*, 45(8):1891–1895, 2009.

[18] L. Ljung. *System Identification: Theory for the User.* PTR Prentice Hall, Upper Saddle River, NJ, USA, 2nd ed. edition, 1999.

[19] P. Lopes dos Santos, T.-P. Azevedo-Perdicoúlis, J.A. Ramos, J.L. Martins de Carvalho, G. Jank, and J. Milhinhos. An LPV modeling and identification approach to leakage detection in high pressure natural gas transportation networks. *IEEE Trans. Control Systems Technology*, 19(1):77–92, 2011.

[20] Riccardo Marino and Patrizio Tomei. *Nonlinear control design.* Information and system sciences. Prentice Hall, London, New York, 1995.

[21] Raman K. Mehra. On the identification of variances and adaptive Kalman filtering. *IEEE Transactions on Automatic Control*, 15(2):175–184, 1970.

[22] R.J. Patton, P.M. Frank, and R. Clarke, editors. *Issues of Fault Diagnosis for Dynamic Systems.* Springer, London, 2000.

[23] Alexandru Ţiclea and Gildas Besançon. Adaptive observer for discrete time state affine systems. In *16th IFAC Symposium on System Identification (SYSID)*, Bruxelles, Belgium, 2012.

[24] Alexandru Ţiclea and Gildas Besançon. Adaptive observer design for discrete time LTV systems. *International Journal of Control*, 89(12):2385–2395, 2016.

[25] R. Tóth, J. Willems, P. Heuberger, and P. Van den Hof. The behavioral approach to linear parameter-varying systems. *IEEE Trans. Automatic Control*, 56(11):2499–2514, 2011.

[26] Ivan Y. Tyukin, Erik Steur, Henk Nijmeijer, and Cees van Leeuwen. Adaptive observers and parameter estimation for a class of systems nonlinear in the parameters. *Automatica*, 49(8):2409–2423, 2013.

[27] Aiping Xu and Qinghua Zhang. Nonlinear system fault diagnosis based on adaptive estimation. *Automatica*, 40(7):1181–1193, 2004.

[28] Qinghua Zhang. Adaptive observer for multiple-input-multiple-output (MIMO) linear time varying systems. *IEEE Trans. on Automatic Control*, 47(3):525–529, 2002.

[29] Qinghua Zhang. Stochastic hybrid system actuator fault diagnosis by adaptive estimation. In *9th IFAC Symposium on Fault Detection, Supervision and Safety of Technical Processes (SAFEPROCESS)*, Paris, September 2015.

[30] Qinghua Zhang. Adaptive Kalman filter for actuator fault diagnosis. In *IFAC World Congress*, page To appear, Toulouse, 2017.

[31] Qinghua Zhang and Michèle Basseville. Statistical detection and isolation of additive faults in linear time-varying systems. *Automatica*, 50(10):2527–2538, 2014.